

Experience Transfer for Multimodal LLM Agents in Minecraft Game

Supplementary Material

6. Appendix

6.1. Algorithmic Details

ICAL Prompt Template. We now describe the ICAL prompting protocol that enables analogy-driven planning over the structured CSD memory. We use a fixed textual template with three parts: (1) system instruction, (2) retrieved exemplars, and (3) the current query.

System instruction. We provide the model with a concise role description and the five transfer dimensions:

```
[SYSTEM]
You are a planning agent in Minecraft.
You receive past tasks described by Contextual State
  Descriptors (CSD)
along five explicit transfer dimensions:

- STRUCTURAL: how entities are spatially arranged
  and organized.
- ATTRIBUTE: visual and physical properties of
  entities.
- PROCEDURAL: the sequence of state transitions and
  actions.
- FUNCTIONAL: what entities are used for and what
  roles they play.
- INTERACTION: how the agent interacts with the
  world and tools.

Your goal is to:
1) Reason by analogy across these five dimensions,
2) Propose new actionable plans for the current state,
  and
3) Output an executable action sequence that follows
  the required schema.
Follow the output format strictly.
```

Retrieved exemplars. For each retrieved memory item $m_{i_k} \in S_K$, we insert a structured exemplar block:

```
[EXAMPLE k]

[GOAL]
{Natural language description of the past task}

[CSD-STRUCT]
{struct_{i_k}}

[CSD-ATTR]
{attr_{i_k}}

[CSD-PROC]
{proc_{i_k}}

[CSD-FUNC]
{func_{i_k}}

[CSD-INTER]
{inter_{i_k}}

[SUCCESSFUL ACTION SEQUENCE]
1. {primitive or high-level action}
2. ...
N. {primitive or high-level action}

[END EXAMPLE k]
```

Current query and instructions. After all exemplars, we append the current CSD and explicit instructions:

```
[CURRENT STATE CSD]

[CSD-STRUCT]
{hat_struct_t}

[CSD-ATTR]
{hat_attr_t}

[CSD-PROC]
{hat_proc_t}

[CSD-FUNC]
{hat_func_t}

[CSD-INTER]
{hat_inter_t}

[INSTRUCTIONS]

1. Compare the current CSD with each EXAMPLE along the
  STRUCTURAL, ATTRIBUTE, PROCEDURAL, FUNCTIONAL,
  and INTERACTION dimensions.
2. Identify transferable patterns (e.g., same recipe
  shape, same
  material family, same process structure, same tool
  function).
3. By analogy, propose ONE new high-level task that is
  feasible
  in the current world state.
4. For this task, output:
  - A short TASK_DESCRIPTION.
  - A step-by-step ACTION_SEQUENCE that can be
  executed.

[OUTPUT FORMAT]

Return a JSON object with the following schema:

{
  "task_description": "...",
  "action_sequence": [
    "step 1: ...",
    "step 2: ...",
    "...",
    "step M: ..."
  ],
  "self_checks": [
    "check 1: ...",
    "check 2: ..."
  ]
}

Only output the JSON object.
```

This template makes the five transfer axes explicit in the context and forces the model to (1) read CSD fields as structured evidence, (2) perform analogy-based reasoning over multiple exemplars, and (3) emit plans and self-checks in a machine-parsable format, which are then consumed by the verifier and executor in the main loop (Figure 6).